

AD-A147 629

TAIL BEHAVIOUR FOR SUPREMA OF EMPIRICAL PROCESSES(U)  
TECHNION - ISRAEL INST OF TECH HAIFA FACULTY OF  
INDUSTRIAL AN. R J ADLER ET AL. SEP 84

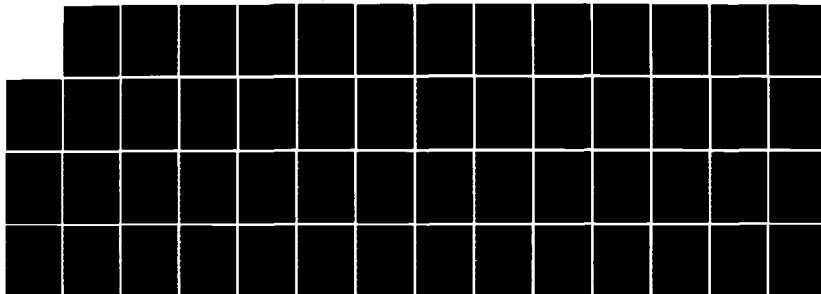
1/1

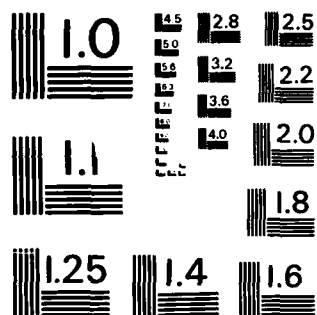
UNCLASSIFIED

AFOSR-TR-84-0937 AFOSR-83-0068

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR. 84-0937

(4)

TAIL BEHAVIOUR FOR SUPREMA OF EMPIRICAL PROCESSES

by

Robert J. Adler and Lawrence D. Brown

Faculty of Industrial Engineering  
Technion - Israel Institute of Technology  
Haifa, Israel

and

Department of Mathematics  
Cornell University  
Ithaca, New York

Grant No. AFOSR 83-0068 , 84-0104

September 13, 1984

Approved for public release:

Distribution unlimited

Prepared for:

Technion Research and Development Foundation Ltd.

P. O. Box 4910

Haifa, Israel

and

European Office of Aerospace Research and Development  
London, England

DTIC  
REFLECTE  
NOV 28 1984  
A

AD A147629

DTIC FILE COPY

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR. 84-0937	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)		7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research	
6a. NAME OF PERFORMING ORGANIZATION Technion Research and Development Foundation	6b. OFFICE SYMBOL (If applicable)	7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical and Information Sciences, Bolling AFB DC 20332	
6c. ADDRESS (City, State and ZIP Code) P.O. Box 4910 Haifa, ISRAEL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-83-0068		
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable) NM	10. SOURCE OF FUNDING NOS	
8c. ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332		PROGRAM ELEMENT NO 61102F	PROJECT NO 2304
		TASK NO A5	WORK UNIT NO
11. TITLE (Include Security Classification) TAIL BEHAVIOUR FOR SUPREMA OF EMPIRICAL PROCESSES			
12. PERSONAL AUTHOR(S) Robert J. Adler and Lawrence D. Brown			
13a. TYPE OF REPORT Technical	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr. Mo. Day) SEP 84	15. PAGE COUNT 50
16. SUPPLEMENTARY NOTATION			
17. COSAT CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB GR	
		Tail behaviour of suprema; empirical processes; Kolmogorow-Smirnov tests; Gaussian random fields.	
19. ABSTRACT (Continue on reverse if necessary and identify by block number)			
<p>We consider multi-variate empirical processes <math>X_n(t) = \sqrt{n}(F_n(t) - F(t))</math>, where <math>F_n</math> is an empirical distribution function based on i.i.d. variables with distribution function <math>F</math>, and <math>t \in \mathbb{R}^k</math>. For <math>X_F</math> the weak limit of <math>X_n</math>, it is shown that</p>			
(CONTINUED)			
20. DISTRIBUTION AVAILABILITY OF ABSTRACT UNCLASSIFIED UNL. MITED <input checked="" type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL MAJ Brian W. Woodruff		22b. TELEPHONE NUMBER (Include Area Code) 707-5027	22c. OFFICE SYMBOL

DD FORM 1473, 83 APR

EDITION OF 1 JAN 73 IS OBSOLETE

84 11 16 017

SECURITY CLASSIFICATION

ITEM #19, ABSTRACT, CONTINUED:

$$c(F,k)\lambda^{2(k-1)}e^{-2\lambda^2} \leq P\{\sup_t X_F(t) > \lambda\} \leq C(k)\lambda^{2(k-1)}e^{-2\lambda^2}$$

for large  $\lambda$  and appropriate constants  $c, C$ . When  $k = 2$  these constants can be identified, thus permitting the development of Kolmogorov-Smirnov tests for bivariate problems. For general  $k$  the bound can be used to obtain sharp upper-lower class results for the growth of  $\sup_t X_n(t)$  with  $n$ .

## SUMMARY

### TAIL BEHAVIOUR FOR SUPREMA OF EMPIRICAL PROCESSES

We consider multi-variate empirical processes  $X_n(t) := \sqrt{n}(F_n(t) - F(t))$ , where  $F_n$  is an empirical distribution function based on i.i.d. variables with distribution functions  $F$ , and  $t \in R^k$ . For  $X_F$  the weak limit of  $X_n$ , it is shown that

$$c(F,k)\lambda^{2(k-1)}e^{-2\lambda^2} \leq P\{\sup_t X_F(t) > \lambda\} \leq C(k)\lambda^{2(k-1)}e^{-2\lambda^2}$$

for large  $\lambda$  and appropriate constants  $c, C$ . When  $k = 2$  these constants can be identified, thus permitting the development of Kolmogorov-Smirnov tests for bivariate problems. For general  $k$  the bound can be used to obtain sharp upper-lower class results for the growth of  $\sup_t X_n(t)$  with  $n$ .



Accession For	
NTIS GRA&I	
DTIC TAB	
Unannounced	
Justification	
By	
Prescription/	
Availability Codes	
Avail and/or	
Unannounced	
A-1	

Running head: Suprema of empirical processes.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)  
NOTICE OF REPRODUCTION  
This technical report is the property of the Air Force Office of Scientific Research.  
Approval for reproduction is required.  
Distribution is limited.  
MATTHEW J. KRAMER  
Chief, Technical Information Division

## 1. INTRODUCTION

It is well known that the Kolmogorov-Smirnov (KS) statistic, based on a sample from any univariate random variable with continuous distribution function (d.f.), is distribution free. It is also well known that in the multivariate situation this is not the case, and it is to this situation that we shall soon direct our efforts. In the beginning, however, Kolmogorov (1933) showed that the one-sided statistic,  $T_n = \sup\{\sqrt{n}(F_n(x) - F(x)) : x \in \mathbb{R}^1\}$ , where  $F$  denotes the underlying d.f. and  $F_n$  the empirical d.f., satisfies

$$(1.1) \quad P\{T_n > \lambda\} \rightarrow e^{-2\lambda^2} \quad \forall \lambda, \text{ as } n \rightarrow \infty.$$

Smirnov (1944) extended this result to the two-sample problem, Feller (1948) gave it a neater proof, and Doob (1949) followed by Donsker (1951, 52) and the theory of weak convergence explained it in terms of the convergence of  $\sqrt{n}(F_n - F)$  to a limiting Gaussian process whose maximum had the tail distribution  $\exp(-2\lambda^2)$ .

In the multivariate case, there is no simple analogue to (1.1), and the best one can hope to obtain is either a limiting distribution for some specific  $F$ , or bounds that may be valid for a family of  $F$ 's sharing, perhaps, some regularity properties. The first attack on this problem was made by Kiefer and Wolfowitz (1958), who showed that if  $T_n^{(k)}$  is the one-sided KS statistic in  $k$  dimensions, then for some  $\alpha = \alpha(k) > 0$  and  $c = c(k) < \infty$ ,

$$(1.2) \quad P\{T_n^{(k)} > \lambda\} \leq ce^{-\alpha\lambda^2} \quad \forall n, \lambda, F.$$

Despite the fact that this bound is obviously very crude, it did at least suffice to prove the existence of a limiting distribution for  $T_n^{(k)}$  as  $n \rightarrow \infty$ . (The full weak convergence of the empirical d.f. to an appropriate limiting Gaussian random field was later established by Dudley (1966, 67).) However, although Kiefer and Wolfowitz established the existence of this limiting distribution, no explicit form for it is known. Indeed, there is only one non-trivial case where reasonably accurate bounds are known, this being the case where  $F$  is uniform on the unit square. Here the limiting distribution of  $\sqrt{n}(F_n - F)$  is that of a pinned Brownian sheet, and fairly close lower and upper bounds on the distribution of its maximum appear in Goodman (1976) and Cabana and Wschebor (1982), respectively. We shall have more to say on this later, when Goodman's lower bound is extended to arbitrary dimensions.

In a classic paper, Kiefer (1961) greatly improved on (1.2) and showed that for all  $\epsilon > 0$  there is a  $c = c(k, \epsilon)$  such that

$$(1.3) \quad P\{T_n^{(k)} > \lambda\} \leq ce^{-2(1-\epsilon)\lambda^2} \quad \forall n, \lambda, F.$$

This is a particularly interesting bound since, viewed as a result on the maximum of the limiting Gaussian field, rather than as a result on  $T_n^{(k)}$  itself, it is one of the few fore-runners of the general inequality for continuous Gaussian processes,  $X(t)$ , that states that for all sufficiently large  $\lambda$

$$(1.4) \quad P\{\sup_t X(t) > \lambda\} \leq e^{-a\lambda^2}, \quad \forall a < (2\sigma^2)^{-1},$$



where  $\sigma^2 = \sup_t \{\text{var } X(t)\}$ . (Fernique (1970, 71), Landau and Shepp (1971), Marcus and Shepp (1971).) Note that since  $\sup_x \{\text{var}[\sqrt{n}(F_n(x) - F(x))]\} = 1/4$ , Kiefer's bound is, today, a simple consequence of (1.4) and weak convergence. Nevertheless, in its time, Kiefer's result was of substantial interest, since it was, to the best of our knowledge, the first time that a uniform bound was placed on the maxima of a large family of Gaussian processes. (The statistical significance of such a lower bound is that it permits construction of "confidence intervals" for an unknown  $F$ .) Furthermore, Kiefer exploited (1.3) to prove a law of the iterated logarithm (LIL) for the multivariate KS statistic.

The main thrust of the current work will be to further refine (1.3), in two directions, and then to investigate the consequences of the refinement. For a start, we shall show (Section 4) that (1.3) can be replaced by: There is a  $c = c(k)$  such that

$$(1.5) \quad P\{T_n^{(k)} > \lambda\} \leq c\lambda^{2(k-1)} e^{-2\lambda^2}, \quad \forall n, \lambda, F.$$

This, as with Kiefer's result, is of interest beyond the KS situation, since, in the Gaussian process setting, it provides a family of processes for which (1.4) can be improved upon. However, we can do better than just (1.5), and we shall also show that as long as  $F$  satisfies mild regularity conditions, there is a  $c = c(F)$  such that

$$(1.6) \quad P\{T_n^{(k)} > \lambda\} \geq c\lambda^{2(k-1)} e^{-2\lambda^2}, \quad \forall n, \lambda.$$

The upper and lower bound together enable us to improve on Kiefer's LIL, and to obtain an exact upper-lower class result in its place (Section 5).

An upper bound similar in spirit to (1.5) has recently been obtained by Alexander (1983). Treating a more general situation of empirical measures,  $v_n$ , indexed by a Vapnik-Červonenkis class of functions,  $F$  say, he showed that

$$P\{\sup_{f \in F} |v_n(f)| > \lambda\} \leq 16\lambda^{2^{1/2}} e^{-2\lambda^2}, \quad \forall \lambda \geq 8,$$

where  $v$  is a strictly positive integer describing the "size" of  $F$ . Alexander's result, while clearly being an improvement on (1.3), also gives, for the cases we consider, an enormous over-estimate of the power of  $\lambda$  in the upper bound.

Unlike Alexander, however, we shall have little to say about the sizes of the constants in our bounds, other than to guarantee their finiteness. Thus, from the point of view of actually applying the KS statistic in a statistical setting, these results are of limited interest. We shall remedy this situation in Section 3, where, for the two-dimensional case, we shall develop an explicit, sharp, upper bound, and a reasonable lower bound. The various applications of these results are spelled out in detail in Brown and Adler (1984). The argument leading to the upper bound is rather interesting, since it is based on finding the worst possible  $F$  (a task actually performed by Kiefer) and comparing it, via Slepian's (1962) inequality, to all other cases. The distribution of the maximum in the worst possible case is what then provides the bound. In fact, this methodology of "comparison" will also be used to obtain the lower bound (1.6), and may, in a certain sense, be considered the main methodological theme of this paper.

The following section is devoted to peripheral and support material. There we obtain lower bounds for the distribution of the maximum of the pinned Brownian sheet in  $k$ -dimensions, and some related distributions. While these do have some intrinsic interest, our main interest in them

will arise from their usefulness as "comparison distributions". We close this section with notation and some background results.

Let  $X_1, X_2, \dots$ , be independent random variables with d.f.  $F(x)$ , which we assume to be continuous, and which can therefore, w.l.o.g., be taken to be concentrated on the unit cube  $I^k = [0,1]^k$  of  $\mathbb{R}^k$  with univariate marginals uniform on  $[0,1]$ . We denote a point in  $I^k$  by either  $x$  or  $(x_1, \dots, x_k)$  and introduce the usual partial order.

$$x < y \iff x_i < y_i, i=1, \dots, k, \quad x, y \in I^k.$$

For  $x < y$  we write  $[x, y]$  for the set  $\prod_{i=1}^k [x_i, y_i]$ , and use  $I_A(\cdot)$  for the indicator function of the set  $A \subset I^k$ . Thus we can formally introduce the empirical d.f.  $F_n$  as

$$(1.7) \quad F_n(x) := n^{-1} \sum_{i=1}^n I_{[0, x]}(X_i).$$

Let  $W_F$  be the pinned Brownian sheet based on  $F$ ; i.e., the zero mean Gaussian process with covariance function

$$(1.8) \quad R_F(x, y) = E\{W_F(x)W_F(y)\} = F(x \wedge y) - F(x)F(y), \quad x, y \in I^k,$$

where  $x \wedge y$  is the coordinatewise minimum  $(x_1 \wedge y_1, \dots, x_k \wedge y_k)$ .

Then, as is well known, ( $k=1$ , Donsker (1952);  $k>1$ , Dudley (1966, 67)),  $\sqrt{n}(F_n - F)$  converges weakly to  $W_F$  in the space of all bounded functions on  $I^k$ . Thus, in particular, if

$$(1.9) \quad T_n^k = T_n^k(F) := \sup\{\sqrt{n}(F_n(x) - F(x)) : x \in I^k\}$$

is the one-sided KS statistic, then

$$(1.10) \quad T_n^k(F) \xrightarrow{L} M_F := \sup\{W_F(x) : x \in I^k\} \quad \text{as } n \rightarrow \infty.$$

This last result provides the obvious motivation for the next two sections, both of which are concerned with the distribution of  $M_F$ . In fact, one can go beyond the central limit result (1.9) to a much stronger embedding type result. However, since we shall not need this result until Section 5, we shall introduce it only then.

## 2. TWO SPECIAL CASES

We consider firstly the distribution of  $M_F$  when  $F$  is the uniform distribution,  $U$  say, on  $I^k$ . We shall, however, require a slightly more general result later, and to this end let  $W^{(k)}$  denote the (unpinned) Brownian sheet on  $I^k$ , i.e., the zero mean Gaussian process with covariance function

$$E\{W^{(k)}(x)W^{(k)}(y)\} = \sum_{i=1}^k (x_i \wedge y_i), \quad x, y \in \mathbb{R}_+^k,$$

and write  $\overset{o}{W}^{(k)}$  for the pinned version of  $W^{(k)}$  on  $I^k$ . Then a version of  $\overset{o}{W}$  can be obtained from  $W$  by the correspondence

$$(2.1) \quad \overset{o}{W}^{(k)}(x) = W^{(k)}(x) - |x|W^{(k)}(1) \quad x \in I^k,$$

where  $|x| = \sum_{i=1}^k x_i$ . In the general notation of the previous section  $\overset{o}{W}^{(k)} \equiv W_U$ . The result we shall need is

### Theorem 2.1

$$(2.2) \quad P\{\sup_{I^k} W^{(k)}(x) > \lambda | W^{(k)}(1) = w\} \geq e^{-2\lambda(\lambda-w)} \sum_{n=0}^{k-1} [2\lambda(\lambda-w)]^n / (n!),$$

a.s., for all  $\lambda > w$ . Furthermore, the case  $w = 0$  yields

$$(2.3) \quad P\{\sup_{I^k} \overset{o}{W}^{(k)}(x) > \lambda\} \geq e^{-2\lambda^2} \sum_{n=0}^{k-1} (2\lambda^2)^n / (n!).$$

When  $k=1$  (2.2) follows immediately from the reflection principle (Feller (1971)). When  $k=2$ , (2.3) is given explicitly in Goodman (1976), and (2.2) is also there implicitly. Cabana and Wschebor (1982) and Park and Skoug (1978) also have (2.2) in the two dimensional case, and

shortly after obtaining the above result for general  $k$ , we received a copy of Cabaña (1982) which states the same result with a virtually identical, albeit more detailed, proof. However, since Cabaña's paper is not readily available, as well as for the sake of completeness, we shall give a brief proof of the theorem.

Proof. The proof proceeds by induction. As noted above, (2.2) is known to be true when  $k=1$  and  $k=2$ . Now write  $a_k(\lambda, w)$  for the conditional probability on the left in (2.2), and define

$$(2.4) \quad h_k(\lambda, w) = P\left\{\sup_{1 \leq k} \left| \frac{W^{(k)}(x) - \lambda}{|x|} \right| > w\right\}.$$

Then, after some calculation, it readily follows from (2.1) that  $a_k(\lambda, w) = h_k(\lambda, -w)$ . If we now follow the formulation of Goodman (1976) of treating the  $(k+1)$ -parameter, real valued,  $W(x)$  as a  $C_0[0,1]^k$  valued, single parameter process, then by applying Goodman's Theorem 2 and mimicking his manipulations on page 980, it is straightforward to establish the relation

$$a_{k+1}(\lambda, w) = \int_{-\lambda}^0 (1 - e^{-2\lambda(w-u)}) h_k(\lambda, du) .$$

Exploiting the above relationship between  $a_k$  and  $h_k$  thus yields a recurrence formula for  $a_{k+1}$ , and it is now a matter of elementary calculus to check the induction hypothesis and so complete the proof.

The second result which, unlike Theorem 2.1, is of little independent interest, will be extremely useful for us later. To state it, we introduce, for each  $\epsilon \in (0, \frac{1}{2})$ , the d.f.  $F_\epsilon(x) = F_\epsilon(x_1, \dots, x_k)$  defined by

$$(2.5) \quad F_{\epsilon}(x) = \begin{cases} \min(x_i) & \text{if } |\min(x_i) - \frac{1}{2}| > \epsilon, \\ \frac{1}{2} - \epsilon + (2\epsilon)^{-k+1} \cdot \prod_{i=1}^k [(x_i - \frac{1}{2} + \epsilon) \wedge 2\epsilon] & \text{otherwise.} \end{cases}$$

Despite its somewhat forbidding appearance,  $F_{\epsilon}$  is a rather simple d.f., distributing total probability  $2\epsilon$  uniformly on the cube  $A_{\epsilon} := [\frac{1}{2}-\epsilon, \frac{1}{2}+\epsilon]^k$ , with the remaining probability of  $1-2\epsilon$  distributed uniformly on that part of the main diagonal of  $I^k$  disjoint from  $A_{\epsilon}$ .

Now let  $\psi_{\epsilon}(u,v)$  denote the two-dimensional normal density with zero means and covariance matrix  $\Sigma_{\epsilon}$  defined by

$$(2.6) \quad \Sigma_{\epsilon} := \begin{bmatrix} (\frac{1}{2}-\epsilon)(\frac{1}{2}+\epsilon) & (\frac{1}{2}-\epsilon)^2 \\ (\frac{1}{2}-\epsilon)^2 & (\frac{1}{2}-\epsilon)(\frac{1}{2}+\epsilon) \end{bmatrix}.$$

Furthermore, let  $\hat{\Sigma}_{\epsilon}$  be the matrix identical to  $\Sigma_{\epsilon}$ , but with the sign of the off-diagonal entries reversed, and let  $\hat{\psi}_{\epsilon}$  be the two-dimensional normal density with zero mean and covariance matrix  $\hat{\Sigma}_{\epsilon}$ . For each positive  $\epsilon$  and  $\lambda$ , and integral  $k$ , set

$$(2.7) \quad \psi_{k,\epsilon}(\lambda) := \int_{-\infty}^{2\epsilon\lambda} \int_{-\infty}^{2\epsilon\lambda} \left\{ \sum_{n=0}^{k-1} \frac{[(u-2\epsilon\lambda)(v-2\epsilon\lambda)/\epsilon]^n}{n!} \right\} \hat{\psi}_{\epsilon}(u,v) du dv,$$

and write  $Q_{\epsilon}(\lambda)$  for the normal quadrant integral

$$(2.8) \quad Q_{\epsilon}(\lambda) := 1 - \int_{-\infty}^{\lambda} \int_{-\infty}^{\lambda} \psi_{\epsilon}(u,v) du dv.$$

We can now state

Theorem 2.2. For every  $\lambda > 0$

$$(2.9) \quad P\{\sup_{I^k} W_{F_\epsilon}(x) > \lambda\} \geq Q_\epsilon(\lambda) + e^{-2\lambda^2} \psi_{k,\epsilon}(\lambda) .$$

In particular, there is a finite  $c = c(\epsilon, k)$  such that for all  $\lambda > 0$

$$(2.10) \quad P\{\sup_{I^k} W_{F_\epsilon}(x) > \lambda\} \geq c \lambda^{2(k-1)} e^{-2\lambda^2} .$$

Proof. We shall obtain only a lower bound for  $P\{\sup_{F_\epsilon}(W_{F_\epsilon}(x) : x \in A_\epsilon) > \lambda\}$ , which, a fortiori, will provide the lower bound required. Let  $a_\epsilon$  and  $b_\epsilon$  be the two extreme corners of  $A_\epsilon$ , i.e.,  $a_\epsilon = (\frac{1}{2} - \epsilon, \dots, \frac{1}{2} - \epsilon)$ ,  $b_\epsilon = (\frac{1}{2} + \epsilon, \dots, \frac{1}{2} + \epsilon)$ . Then define the process  $Z(x)$  on  $I^k$  by

$$Z(x) := (2\epsilon)^{-\frac{1}{2}} W_{F_\epsilon}(a_\epsilon + 2\epsilon x) - (1 - |x|) W_{F_\epsilon}(a_\epsilon) - |x| W_{F_\epsilon}(b_\epsilon)$$

Then it is straightforward to check that  $Z(x)$  is a standard pinned sheet on  $I^k$ , as in (2.1). Consequently, for  $(u, v) < \lambda$ , it follows that

$$(2.11) \quad P\{\sup_{A_\epsilon} W_{F_\epsilon}(x) > \lambda \mid W_{F_\epsilon}(a_\epsilon) = u, W_{F_\epsilon}(b_\epsilon) = v\} \\ = P\{\sup_{I^k} \left( \frac{W^{(k)}(x) - (\lambda - u)/\sqrt{2\epsilon}}{|x|} \right) \geq \frac{u - v}{\sqrt{2\epsilon}}\}$$

But this is precisely the probability defined at (2.4). Thus, using the equivalence noted there between this probability and  $a_k$ , we can bound it by Theorem 2.1. Using this bound, (2.11), and the fact that the joint density of  $\langle W_{F_\epsilon}(a_\epsilon), W_{F_\epsilon}(b_\epsilon) \rangle$  is given by  $\psi_\epsilon$ , we obtain



$$(2.12) \quad P\{\sup_{I^k} W_{F_\epsilon}(x) > \lambda\} \geq P\{W_{F_\epsilon}(a_\epsilon) > \lambda \text{ or } W_{F_\epsilon}(b_\epsilon) > \lambda\} +$$

$$\int_{-\infty}^{\lambda} \int_{-\infty}^{\lambda} e^{-(\lambda-u)(\lambda-v)/\epsilon} \left\{ \sum_{n=0}^{k-1} [(\lambda-u)(\lambda-v)/\epsilon]^n / n! \right\} \psi_\epsilon(u, v) du dv.$$

Consider the integrand, and make the transformations  $x=u-\lambda(1-2\epsilon)$ ,  $y=v-\lambda(1-2\epsilon)$ . Tedious but straightforward algebra yields that it is equivalent to

$$e^{-2\lambda^2\epsilon} \psi_\epsilon(x, y) \sum_{n=0}^{k-1} [(x-2\epsilon\lambda)(y-2\epsilon\lambda)/\epsilon]^n / n!$$

Substituting this into (2.12), changing the bounds on the integral, and replacing the rightmost probability by  $Q_\epsilon(\lambda)$  now yields (2.9), as required.

To obtain (2.10) from (2.9) simply take  $\lambda$  large enough so that the dominant term in the sum in  $\psi_{\epsilon, k}$  is  $O(\lambda^{2(k-1)})$ . Then choose an appropriate  $c$  to make (2.10) work. This completes the proof.

In what follows we shall be primarily interested in the asymptotic lower bound (2.10), which will be used to prove results of theoretical interest. The explicit expression (2.9) has, however, some practical value for statistical hypothesis testing, and this is discussed in Brown and Adler (1984), where the bound is actually tabulated for a number of cases.

In general, we shall use Theorem 2.2 to form a basis for comparison between the maxima of pinned sheets based on different d.f.s. The crucial result that underlies all these comparisons is a basic result of Slepian (1962), which we record here as

Lemma 2.1 (Slepian's inequality). Let  $X$  and  $Y$  be two zero mean Gaussian processes defined over some set  $T$ . If  $\text{var } X(t) = \text{var } Y(t)$ ,  $\forall t \in T$ , and

$$(2.13) \quad \text{cov}(X(t), X(s)) \leq \text{cov}(Y(t), Y(s)) \quad \forall s, t \in T.$$

then

$$(2.14) \quad P\{\sup_T X(t) > \lambda\} \geq P\{\sup_T Y(t) > \lambda\} \quad \forall \lambda.$$

Note that Slepian's inequality does not extend to comparisons of  $|\sup X|$  and  $|\sup Y|$ , and so the sharp results of the following section are not easily extendable to the two-sided KS statistic. Nevertheless, we can always use the fact that for symmetric processes

$$(2.15) \quad P\{\sup X > \lambda\} \leq P\{\sup |X| > \lambda\} \leq 2P\{\sup X > \lambda\}$$

to obtain bounds for the two-sided case. For the bounds of section 4, in which constants are not identified, this is clearly sufficient.

We now consider, as an example of our "comparison methodology" the two-dimensional case.

### 3. THE TWO-DIMENSIONAL CASE.

Throughout this section, we shall denote points in  $I^2$  by  $(x,y)$ , and  $F$  will denote a continuous d.f. on  $I_2$  possessing uniform marginals. The degenerate distribution, uniform on the negative slope diagonal  $x+y=1$  will be denoted by  $G(x,y)$ ; i.e.,

$$(3.1) \quad G(x,y) = (x+y-1)^+, \quad (x,y) \in I^2.$$

Our aim in this section will be to devise good (non-asymptotic) bounds for  $P\{\sup_{I^2} W_F > \lambda\}$ . We start with

Theorem 3.1 For any two-dimensional d.f.  $F$  satisfying the above conditions, and for any  $\lambda > 0$

$$(3.2) \quad P\{\sup_{I^2} W_F(x) > \lambda\} \leq P\{\sup_{I^2} W_G(x) > \lambda\}.$$

Furthermore

$$(3.3) \quad P\{\sup_{I^2} W_G(x) > \lambda\} \leq \sum_{n=1}^{\infty} (8n^2\lambda^2 - 2)e^{-2n^2\lambda^2}.$$

Proof. Let  $m$  be the mapping from  $I^2$  into  $I^2$  defined by

$$(3.4) \quad G(m(x)) = G(m_1(x), m_2(x)) = F(x), \quad \forall x \in I^2$$

$$(3.5) \quad m_2(x) - m_1(x) = x_2 - x_1 \quad \forall x \in I^2.$$

We must check that  $m$  is well defined. For given  $x$ , note that  $m(x)$

lies on the line  $l_x := (x_1 + \mu, x_2 + \mu)$ ,  $\mu \in \mathbb{R}$ , and  $G(x_1 + \mu, x_2 + \mu)$  is clearly non-decreasing in  $\mu$ . Indeed, (3.1) and a little elementary geometry show that  $G(x_1 + \mu, x_2 + \mu)$  is strictly increasing for

$$(3.6) \quad \frac{1}{2}(1-x_1-x_2) \leq \mu \leq 1 - (x_1 \vee x_2)$$

When  $\mu = \frac{1}{2}(1-x_1-x_2)$ , then  $G(x_1+\mu, x_2+\mu) = 0 \leq F(x)$ . When  $\mu = 1-(x_1 \vee x_2)$ , then  $G(x_1+\mu, x_2+\mu) = 1 - |x_1 - x_2|$ . Suppose, w.l.o.g., that  $x_1 \geq x_2$ . Then  $(x_1 + \mu, x_2 + \mu) = (1, x_2 + (1-x_1))$ . Applying these facts, together with the uniformity of the marginals of  $F$  and the natural monotonicity of  $F$ , we obtain

$$G(x_1 + \mu, x_2 + \mu) = F(1, x_2 + (1-x_1)) \geq F(x) .$$

Thus, within the range (3.6) there is, by the continuity and strict monotonicity of  $G$ , exactly one  $\mu$  satisfying  $G(x_1 + \mu, x_2 + \mu) = F(x)$ . Hence the map  $m$  is well defined.

Now consider the processes  $W_F$  and  $W_G$ . We shall compare  $\sup\{W_F(x) : x \in I^2\}$  to  $\sup\{W_G(x) : x \in m(I^2)\} \leq \sup\{W_G(x) : x \in I^2\}$ . Note firstly that for  $x \in I^2$

$$(3.7) \quad \text{var } W_F(x) = \text{var } W_G(m(x)) ,$$

a simple consequence of (3.4) and (1.8). Consider

$$(3.8) \quad \text{cov}(W_F(x), W_F(y)) = F(x \wedge y) - F(x)F(y) .$$

Suppose  $x \wedge y = x$ . Then

$$F(x \wedge y) = F(x) = G(m(x)) \geq G(m(x) \wedge m(y))$$

Thus, in this case (as in the analogous case  $x \wedge y = y$ )

$$(3.9) \quad \text{cov}(W_F(x), W_F(y)) \geq \text{cov}(W_G(m(x)), W_G(m(y))).$$

If we can also establish (3.9) in general, then we shall have completed the proof of the first part of the theorem, viz. (3.2), since (3.7) and (3.9) are precisely the ingredients for Slepian's inequality.

Thus, consider (3.8) for  $x, y$  with  $x_1 > y_1$  and  $x_2 < y_2$ . (The remaining case is handled analogously.) Then  $x \wedge y = (y_1, x_2)$ . Write  $w = (m_1(y), m_2(x))$ . There are three possible cases to consider:  $m(x) \geq w \geq m(y)$ ,  $m(y) \geq w \geq m(x)$ ,  $w = m(x) \wedge m(y)$ . We shall consider only the third case explicitly, but the reasoning is valid for all the cases. Note (drawing a picture helps to see the inequalities) that

$$\begin{aligned} F(x \wedge y) &= F(y_1, x_2) \\ &\geq [F(x) - (x_1 - y_1)] \vee [F(y) - (y_2 - x_2)] \quad \text{by marginal uniformity} \\ &\geq \frac{1}{2} \{ F(x) + F(y) - [(x_1 - x_2) - (y_1 - y_2)] \} \\ &= \frac{1}{2} \{ G(m(x)) + G(m(y)) - [(m_1(x) - m_2(x)) - (m_1(y) - m_2(y))] \} \\ &\quad \text{by (3.4), (3.5)} \\ &\geq m_2(x) + m_1(y) - 1. \quad \text{by (3.1).} \end{aligned}$$

Hence, if  $m_2(x) + m_1(y) - 1 \geq 0$  then the above yields

$$(3.10) \quad F(x \wedge y) \geq G(m_1(y), m_2(x)).$$

On the other hand, if  $m_2(x) + m_1(y) - 1 < 0$ , then  $G(m_1(y), m_2(x)) = 0$  and so (3.10) is trivially true. Thus, in general,

$$F(x,y) \geq G(m_1(y), m_2(x)) = G(w) = G(m(x) \wedge m(y)) .$$

From this we immediately obtain (3.9) and the proof of (3.2).

It remains to establish the inequality (3.3). To this end, let  $\overset{\circ}{W}(t)$ ,  $t \in [0,1]$ , be a standard Brownian bridge with covariance function

$$(3.11) \quad E\{\overset{\circ}{W}(t)\overset{\circ}{W}(s)\} = (sat) - st .$$

Define the two-parameter field  $X$  on  $I^2$  by

$$X(x_1, x_2) = \begin{cases} \overset{\circ}{W}(x_1) - \overset{\circ}{W}(1-x_2) & x_1 + x_2 - 1 \geq 0 \\ 0 & x_1 + x_2 - 1 \leq 0 \end{cases}$$

Then comparison of covariance functions shows that  $X$  is a version of  $W_G$ . Thus

$$\begin{aligned} (3.12) \quad P\{\sup_G \overset{\circ}{W}(x_1, x_2) > \lambda\} &= P\{\sup[\overset{\circ}{W}(x_1) - \overset{\circ}{W}(1-x_2) : x_1 + x_2 - 1 \geq 0] > \lambda\} \\ &= P\{\sup[\overset{\circ}{W}(s) - \overset{\circ}{W}(t) : s > t] > \lambda\} \\ &\leq P\{\sup[\overset{\circ}{W}(s) - \overset{\circ}{W}(t) : s, t \in [0,1]] > \lambda\} \\ &\leq P\{[\sup_{[0,1]} \overset{\circ}{W}(s)]^+ + \sup_{[0,1]} \overset{\circ}{W}(s)]^- > \lambda\} \end{aligned}$$

But the last probability is known exactly, having been determined in Kac, Kiefer and Wolfowitz (1955, equation (4.6)), and is precisely the sum given on the right of (3.3), and so we are finished.

Remark. Note that the two inequalities following (3.12) are far from sharp, and a little reflection shows that each inequality, while retaining a bound of the right order of magnitude, "costs", roughly, a factor of two, i.e., we expect that the final upper bound is too large by a factor of four. Indeed, comparison of the general upper bound (3.3) with the specific lower bound in the uniform case, (2.3) with  $k=2$ , shows, for large  $\lambda$ , a difference between the bounds of precisely a factor of four. Clearly, a much better upper bound than (3.3) is given by  $P\{\sup[\hat{W}(s) - \hat{W}(t): 0 \leq t < s \leq 1] \geq \lambda\}$ , (c.f. (3.12)), but this seems hard to calculate. However, numerical estimates of this probability are easy to obtain via simulation, and some are listed in Brown and Adler (1984). Furthermore, calculation of (3.3) and comparison with (2.3) for moderate  $\lambda$ , say  $\lambda \in [1, 3]$ , yields that (3.3) overestimates the true probability by less than a factor of four, and that the KS test statistics derivable from (3.3) are in fact quite useful. For details see Brown and Adler (1984).

We now turn to the more difficult problem of finding a uniform lower bound for the two-dimensional case. Here we shall need to impose assumptions on  $F$  in order to avoid degeneracies. (e.g.,  $F$  concentrated on the diagonal  $x_1=x_2$ , which reduces to the one-dimensional case.) Let  $\|x\| = |x_1| + |x_2|$  denote the "city block" norm of  $X$ . Then we shall prove

Theorem 3.2 Let  $F$  be a d.f. on  $I^2$ , with uniform marginals, such that there exists an  $x_0 \in I^2$ , a neighbourhood  $N$  of  $x_0$ , and a constant  $\beta \in (0, 1]$  satisfying

$$(3.13) \quad F(x_0) = \frac{1}{2}$$

and

$$(3.14) \quad \text{If } x, y \in N \text{ and either } x_1 = y_1 \text{ or } x_2 = y_2 \text{ then } |F(x) - F(y)| > \beta ||x - y||.$$

Then there exists a finite  $c = c(F) > 0$  such that

$$(3.15) \quad P\{\sup W_F(x) > \lambda\} \leq c\lambda^2 e^{-2\lambda^2}.$$

Remarks. Theorem 3.2, as it stands, is a special case of the more general result Theorem 4.2. What makes it of special interest, however, is the fact that in two dimensions it is possible to obtain estimates for  $c$ . We shall discuss these at the end of the proof. Furthermore, the two dimensional case turns out to be somewhat simpler than its higher dimensional analogue, thereby making its proof more transparent and interesting.

It is clear that the conditions of Theorem 3.2 hold if  $F$  has a density bounded away from zero. However, absolute continuity is not a requisite of the theorem, and it is easy to build examples of non-absolutely continuous  $F$  satisfying (3.13) and (3.14). A trivial example is the extremal case, (3.1).

Proof of Theorem 3.2. The aim of the proof will be to compare  $W_F$  with  $W_{F_\epsilon}$ , where  $F_\epsilon$  is the distribution function (2.5) of the preceding section, and then use Slepian's inequality and Theorem 2.2 to complete the argument. The comparison will only be possible over a region in the neighbourhood of  $(\frac{1}{2}, \frac{1}{2})$  in the domain of  $W_{F_\epsilon}$ , together with a subset of  $N$  in the domain of  $W_F$ , but it will turn out that such a comparison will suffice for our purposes. We start by



building a mapping between the above two neighbourhoods, and by noting that the reader's path through the forthcoming algebra will be considerably simpler if he follows the argument graphically with pen and paper.

For  $x \in \mathbb{R}^2$ , let  $\hat{x}$  be the projection of  $x$  on the diagonal  $\{x: x_1 = x_2\}$ , i.e.,  $\hat{x}$  has both coordinates equal to  $\frac{1}{2}(x_1 + x_2)$ . Define

$$(3.16) \quad d = d(N) := \inf\{\|x - x_0\|, x \notin N\}.$$

Let  $\epsilon = d\beta/3$  and define a map  $m$  from  $A_\epsilon$  into  $N$  satisfying

$$(3.17) \quad m(y) - \hat{m}(y) = x_0 - \hat{x}_0 + (y - \hat{y})/\beta.$$

and

$$(3.18) \quad F_\epsilon(y) = F(m(y)),$$

where  $F_\epsilon$  is defined at (2.5). It is necessary to demonstrate that this map is well defined and one-one.

To this end, fix  $y \in A_\epsilon$  and let  $m_0 = x_0 + (y - \hat{y})/\beta$ . Also, let  $m_\mu = m_0 + \mu(1,1)$ , for real  $\mu$ . Now note that, by (3.14),  $F(m_\mu)$  is strictly increasing in  $\mu$  as long as  $m_\mu \in N$ . Furthermore, if  $\mu = \epsilon/\beta$ , then  $m_\mu \geq x_0$ , since  $(y - \hat{y})_1 \leq \epsilon$  for  $y \in A_\epsilon$ . Similarly,  $\mu = -\epsilon/\beta$  implies  $m_\mu < x_0$ . Consequently

$$(3.19) \quad F(m_{-\epsilon/\beta}) \leq \frac{1}{2} \leq F(m_{\epsilon/\beta}).$$

Now consider for what values of  $\mu$  we shall have  $m_\mu \in N$ .

Clearly  $(m_\mu)_i - (x_0)_i < \epsilon/\beta + \mu$ ,  $i=1,2$ , since  $y \in A_\epsilon$ . Hence  $m_\mu \in N$  for  $|u| < 2\epsilon/\beta$ . Now take  $\mu > \mu'$ , with  $|\mu|, |\mu'| < 2\epsilon/\beta$ . Then, by (3.14),

$$F(m_\mu) - F(m_{\mu'}) > \beta(\mu - \mu').$$

But now it follows that there is a unique  $\mu \in [-2\varepsilon/\beta, 2\varepsilon/\beta]$  such that  $F(m_\mu) = F_\varepsilon(y)$ , since  $y \in A_\varepsilon$  implies  $\frac{1}{2} - \varepsilon \leq F_\varepsilon(y) \leq \frac{1}{2} + \varepsilon$ . Let  $m(y) = m_\mu$  for this  $\mu$ . Then clearly (3.18) is satisfied, as is (3.17), so that  $m$  is well defined for each  $y \in A_\varepsilon$ . Furthermore, the above argument also establishes that  $m$  is one-one. This completes the first part of the argument.

Let  $m(A_\varepsilon)$  be the image of  $A_\varepsilon$  under the mapping  $m$ , and consider  $W_F(x)$  for  $x \in m(A_\varepsilon)$ . Clearly, for  $y \in A_\varepsilon$ , (3.18) and (1.8) imply

$$(3.20) \quad E\{W_F^2(m(y))\} = E\{W_F^2(y)\}.$$

Now take  $y_1 < y_2$ ,  $y_1, y_2 \in A_\varepsilon$ . Then  $F_\varepsilon(y_1) < F_\varepsilon(y_2)$  and so  $F(m(y_1)) < F(m(y_2))$ . Thus (1.8) immediately yields

$$\begin{aligned} (3.21) \quad E\{W_F(m(y_1))W_F(m(y_2))\} &= F(m(y_1) \wedge m(y_2)) - F(m(y_1))F(m(y_2)) \\ &\leq F(m(y_1)) \wedge F(m(y_2)) - F(m(y_1))F(m(y_2)) \\ &= F_\varepsilon(y_1) - F_\varepsilon(y_1)F_\varepsilon(y_2) \\ &= E\{W_{F_\varepsilon}(y_1)W_{F_\varepsilon}(y_2)\}. \end{aligned}$$

By symmetry, (3.21) also holds for  $y_1 > y_2$ . Now suppose  $y_3 = y_1 \wedge y_2$  is distinct from both  $y_1$  and  $y_2$ . Set  $x_i = m(y_i)$ ,  $i=1,2$ , and  $x_3 = x_1 \wedge x_2$ . Observe, either geometrically or algebraically, that

$$(3.22) \quad ||y_1 - y_3|| + ||y_2 - y_3|| = ||(y_1 - \hat{y}_1) - (y_2 - \hat{y}_2)||.$$

Thus, since  $F_\varepsilon$  has uniform marginals, and  $y_1 \in A_\varepsilon$ ,

$$\begin{aligned}
 (3.23) \quad & F_\epsilon(y_1) + F_\epsilon(y_2) - 2F_\epsilon(y_3) \\
 & \leq (\frac{1}{2} + \epsilon)(||y_1 - y_3|| + ||y_2 - y_3||) \\
 & = (\frac{1}{2} + \epsilon)||y_1 - \hat{y}_1 - (y_2 - \hat{y}_2)||.
 \end{aligned}$$

Now suppose that  $x_2 \geq x_1$ . We shall show that this is impossible.

Write  $x_4 = ((x_1)_1, (x_2)_2)$ . Then, by geometry and assumption (3.14)

$$\begin{aligned}
 F(x_2) - F(x_1) &= F(x_2) - F(x_4) + F(x_4) - F(x_1) \\
 &\geq \beta(||x_2 - x_4|| + ||x_4 - x_1||) \\
 &\geq \beta||x_2 - x_1 - (\hat{x}_2 - \hat{x}_1)|| \\
 &= ||(y_2 - y_1) - (\hat{y}_2 - \hat{y}_1)||,
 \end{aligned}$$

the last line following from (3.17). The above and (3.23) now yield

$$\begin{aligned}
 0 < F(x_2) - F(x_1) &= F_\epsilon(y_2) - F_\epsilon(y_1) \\
 &\leq F_\epsilon(y_2) + F_\epsilon(y_1) - 2F_\epsilon(y_3) \\
 &\leq (\frac{1}{2} + \epsilon)\{F(x_2) - F(x_1)\},
 \end{aligned}$$

which, since  $\epsilon < \frac{1}{2}$ , is clearly untenable. Thus we cannot have  $x_2 \geq x_1$ , nor, by symmetry,  $x_1 \geq x_2$ . Consequently,  $x_3 = x_1 \wedge x_2$  is distinct from both  $x_1$  and  $x_2$ . Then, again by geometry, assumption, and (3.22), we have

$$\begin{aligned}
 F(x_1) + F(x_2) - 2F(x_3) &\geq \beta(||x_1 - x_3|| + ||x_2 - x_3||) \\
 &= \beta||x_1 - \hat{x}_1 - (x_2 - \hat{x}_2)|| \\
 &= ||(y_1 - \hat{y}_1) - (y_2 - \hat{y}_2)|| \\
 &= ||y_1 - \hat{y}_1|| + ||y_2 - \hat{y}_2|| \\
 &\geq (\frac{1}{2} + \epsilon)^{-\frac{1}{2}}\{F_\epsilon(y_1) + F_\epsilon(y_2) - 2F_\epsilon(y_3)\} \\
 &> F_\epsilon(y_1) + F_\epsilon(y_2) - 2F_\epsilon(y_3).
 \end{aligned}$$

Thus, since  $F(x_i) = F_\epsilon(y_i)$ ,  $i=1,2$ ,

$$F_\epsilon(y_3) = F(m(y_3)) > F(x_3) .$$

From this it immediately follows that for all  $y_1, y_2 \in A_\epsilon$

$$E\{W_{F_\epsilon}(y_1)W_{F_\epsilon}(y_2)\} \geq E\{W_F(m(y_1))W_F(m(y_2))\} ,$$

with strict inequality if  $x_1 \wedge x_2 \neq x_i$ ,  $i=1,2$ . But this is all we need, for by Slepian's inequality,

$$\begin{aligned} P\{\sup_{I^2} W_F(x) > \lambda\} &\geq P\{\sup_{m(A_\epsilon)} W_F(x) > \lambda\} \\ &\geq P\{\sup_{A_\epsilon} W_{F_\epsilon}(x) > \lambda\}. \end{aligned}$$

The last probability is precisely that given by the RHS of (2.9), which, as we have already noted, is asymptotically of the form  $c\lambda^2 e^{-2\lambda^2}$ . This completes the proof of the theorem.

We close this section with two remarks. The first on the constant  $c$  of Theorem 3.2, or, to be more precise, on an exact lower bound for  $P\{\sup W_F(x) > \lambda\}$ . It is clear from the argument that such a bound is given by  $Q_\epsilon(\lambda) + \psi_{2,\epsilon}(\lambda)$ , with  $\epsilon = d\beta/3$ . (c.f. (2.7), (2.8), (3.16).) If we consider the case of  $F$  uniform, the optimal choice of  $d\beta$ , so as to maximise  $\epsilon$ , is  $d = \beta = 1/(2\sqrt{2})$ , yielding  $\epsilon \approx 0.04$ . This is, of course, much smaller than the  $\epsilon = \frac{1}{2}$  that a sharp argument would give. Nevertheless, the numerical consequences of this lack of sharpness are not quite as bad as one might imagine. For details, see Brown and Adler (1984).

It is interesting to note that there are " $1\frac{1}{2}$ -dimensional" d.f.'s that yield supremum tail probabilities strictly between the one-dimensional  $O(e^{-2\lambda^2})$  and two-dimensional  $O(\lambda^2 e^{-2\lambda^2})$ . As an example, take  $H$  to be the d.f. on  $I^2$  with density

$$(3.24) \quad h(x,y) = \begin{cases} 2 & (x,y) < (\frac{1}{2}, \frac{1}{2}) \text{ or } (x,y) > (\frac{1}{2}, \frac{1}{2}) \\ 0 & \text{otherwise.} \end{cases}$$

Clearly,  $H$  fails to satisfy the conditions of Theorem 3.2. However, it is a relatively easy exercise to estimate the exceedence probabilities of  $W_H$ , using the fact that the two processes

$$W_1(x,y) := \sqrt{2}\{W_H(x/2, y/2) - W_H(\frac{1}{2}, \frac{1}{2})\}$$

$$W_2(x,y) := \sqrt{2}\{W_H(1-x/2, 1-y/2) - W_H(\frac{1}{2}, \frac{1}{2})\},$$

$(x,y) \in I^2$ , are both versions of the pinned Brownian sheet  $W^{(2)}$ .

This fact, together with Theorems 2.1 and 3.2, conditioning on and then integrating out  $W_H(\frac{1}{2}, \frac{1}{2})$ , readily yields

$$(3.25) \quad P\{\sup_{I^2} W_H(x) > \lambda\} = O(\lambda e^{-2\lambda^2}),$$

thus indicating that non-even powers of  $\lambda$  in tail bounds cannot be excluded. (Indeed, there is no good reason even to exclude non-integer powers, as these do occur as tail bounds for other classes of Gaussian processes; see, for example, Section 12.2 of Leadbetter, Lindgren and Rootzen (1983).)

#### 4. BOUNDS FOR THE GENERAL CASE

Our aim in this section will be to obtain, in  $k > 2$  dimensions, bounds of the same general form as those we have just obtained for two dimensions. In particular, if  $F$  is a continuous d.f. on  $I^k$  with uniform (one-dimensional) marginals, then the two central results are as follows:

Theorem 4.1 There exist constants  $c_k$ ,  $k \geq 1$ , independent of  $F$  and  $\lambda$ , such that for  $F$  as above

$$(4.1) \quad P\{\sup W_F(x) > \lambda\} \leq c_k \lambda^{2(k-1)} e^{-2\lambda^2}.$$

Theorem 4.2 Suppose, in addition to the above, there exists an  $x_0 \in I^k$ , a neighbourhood  $N$  of  $x_0$ , and a constant  $\psi > 0$  satisfying

$$(4.2) \quad F(x_0) = \frac{1}{2},$$

(4.3) Throughout  $N$ ,  $F$  possesses continuous first order partial derivatives  $\psi_1 := \partial F / \partial x_1$  satisfying

$$\inf_1 \inf_N \psi_1(x) = \psi > 0.$$

Then for each such  $F$  there exists a constant  $c = c(F)$ , independent of  $\lambda$ , such that

$$(4.4) \quad P\{\sup_{I^k} W_F(x) > \lambda\} \geq c \lambda^{2(k-1)} e^{-2\lambda^2}.$$

Both of these results, while clearly indicating the correct order of magnitude behaviour of the tail of  $\sup W_F$ , are considerably weaker

than their two-dimensional counterparts, since the style of their proofs will be such that it will be impossible to closely monitor inequalities so as to estimate the constants of the bounds. Consequently, the statistical value of Theorems 4.1 and 4.2 is somewhat limited. Nevertheless, they have interesting probabilistic consequences, as we shall see in Section 5, as well as being of intrinsic interest for the reasons mentioned in the introduction.

We shall prove Theorem 4.1 first, by a method totally different from that used for the two-dimensional upper bound. There, recall, the argument was based on finding a "worst possible  $F$ ". In dimensions three and above there seems to be no analogous unique worst  $F$ , and the proof is forced to take a different route. We start with some necessary lemmas, for which we define the following event for

$x_1, x_2 \in I^k$ ,  $x_1 \leq x_2$  and  $\lambda > 0$

$$(4.5) \quad A = A(x_1, x_2, \lambda) := \{ \sup(W_F(x) : x_1 \leq x \leq x_2) > \lambda \},$$

Also, write

$$(4.6) \quad \sigma^2(x) := \text{var}(W_F(x)) = F(x)[1-F(x)].$$

Lemma 4.1 Take  $\frac{1}{2} < \alpha < \frac{1}{2}$ ,  $x_1, x_2 \in I^k$ ,  $x_1 \leq x_2$  and  $\lambda > 1$ . If

$$(4.7) \quad \alpha \leq F(x_1) \leq F(x_2) \leq 1-\alpha,$$

and

$$(4.8) \quad F(x_2) - F(x_1) \leq \frac{1}{2} \alpha^2 \lambda^{-2},$$

then

$$(4.9) \quad P\{A\} \leq O(\lambda^{-1} \exp(-\lambda^2/2\sigma^2(x_1))),$$

where, for any function  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,

$$f(a) \leq O(a) \iff \limsup_{a \rightarrow \infty} (f(a)/a) \leq K < \infty.$$

Proof. Since it is generally difficult to work with the maxima of the pinned sheet  $W_F$ , the main idea of the proof is to relate  $W_F$  to its unpinned version,  $Z_F$ , where  $Z_F$  is the zero mean Gaussian field on  $I^k$  satisfying

$$E\{Z_F(x) \cdot Z_F(y)\} = F(x \wedge y) .$$

Then  $Z_F(x) - F(x)Z_F(1)$  is a version of  $W_F$ , so that using this version in all that follows, we can write

$$(4.10) \quad W_F(x) = V(x) - [F(x) - F(x_1)]Z(1) \quad x \in [x_1, x_2]$$

where

$$(4.11) \quad V(x) := W(x_1) + [Z_F(x) - Z_F(x_1)] \quad x \in [x_1, x_2] .$$

The idea of the proof is that for  $\lambda$  large (4.8) implies the second term in (4.10) will be small, while  $V(x)$  will be close to  $W(x_1)$ .

Note firstly, by direct calculation of covariances, that  $W_F(x)$  and  $Z_F(1)$  are independent, so that with  $A$  as at (4.5)

$$(4.12) \quad P\{A\} = 2P\{A \text{ and } Z_F(1) \geq 0\} .$$

Thus, by (4.10)

$$(4.13) \quad \begin{aligned} P\{A\} &\leq 2P\{Z_F(1) \geq 0 \text{ and } \sup_{[x_1, x_2]} V(x) > \lambda\} \\ &\leq 2P\{\sup_{[x_1, x_2]} V(x) > \lambda\} \end{aligned}$$

To bound the last probability, write  $V(x) = W(x_1) + U(x)$ , where

$$U(x) := Z_F(x) - Z_F(x_1)$$

is independent of  $W(x_1)$ . Suppose we can show the existence of a finite



$c > 0$  such that for all  $\eta > 0$

$$(4.14) \quad P\left\{\sup_{[x_1, x_2]} U(x) > \eta\right\} \leq c(1 + \eta\lambda/\alpha)^{-1} e^{-\eta^2\lambda^2/\alpha^2}.$$

Then, allowing  $c$  to vary from line to line, and setting  $\sigma^2 = \sigma^2(x_1)$  for notational convenience, we have

$$\begin{aligned} P\{A\} &\leq 2P\{W(x_1) \geq \lambda\} + \int_{-\infty}^{\lambda} P\{\sup U(x) \geq \lambda - w\} dP\{W(x_1) \leq w\} \\ &\leq c\lambda^{-1} \exp(-\lambda^2/2\sigma^2) + c \int_{-\infty}^{\lambda} (1 + \lambda - w)^{-1} \exp\left\{-\frac{\lambda^2(\lambda-w)^2}{\alpha^2} - \frac{w^2}{2\sigma^2}\right\} dw, \end{aligned}$$

on using standard inequalities for the first probability, and (4.14) for the integrand, after noting  $\lambda > 1$  and  $\alpha < 1/2$ . Standard integration yields that the integral is  $O(\lambda^{-1} e^{-\lambda^2/2\sigma^2})$ . This proves the lemma.

Thus all that remains is to establish (4.14).

A straightforward application of the multivariate "reflection principle" yields

$$P\left\{\sup_{[x_1, x_2]} U(x) > \lambda\right\} \leq 2^k P\{U(x_2) > \lambda\}.$$

By (4.8)  $\text{var}U(x_2) \leq \frac{1}{2} \alpha^2 \lambda^{-2}$ , so that (4.14) now follows by standard inequalities.

Without much extra work we can also prove a stronger version of the preceding lemma. Under the conditions of the lemma, we have, for  $x_1 \leq x \leq x_2$  that

$$\sigma^2(x) \geq \sigma^2 - 5\alpha^2/(4\lambda^2).$$

Consequently

$$\begin{aligned} \frac{\lambda^2}{2\sigma^2(x)} - \frac{\lambda^2}{2\sigma^2} &\leq \frac{\lambda^2}{2[\sigma^2 - 5\alpha^2/4\lambda^2]} - \frac{\lambda^2}{2\sigma^2} \\ &= \frac{5\alpha^2/4}{2\sigma^2[\sigma^2 - 5\alpha^2/4\lambda^2]} = o(1) \end{aligned}$$

Thus lemma 4.1 immediately yields

Lemma 4.2 Under the conditions and notation of Lemma 4.1

$$(4.15) \quad P\{A\} \leq O(\lambda^{-1} \exp(-\lambda^2/2\hat{\sigma}^2)) ,$$

where  $\hat{\sigma}^2 = \inf\{\sigma^2(x) : x_1 \leq x \leq x_2\}$ .

To state the next lemma define the event

$$B = \{\sup(W_F(x) : F(x) \leq \alpha \text{ or } F(x) \geq 1-\alpha) > \lambda\}$$

Lemma 4.3 Let  $\alpha \in (0, 1/2)$ , and  $\beta \in (1, (4\alpha(1-\alpha))^{-1})$ . Then

$$(4.16) \quad P\{B\} \leq O(\exp(-2\beta\lambda^2)) .$$

Proof. This is a straightforward application of (1.4), on noting that  $F(x) < \alpha$  and  $F(x) > 1-\alpha$  both imply  $(2\sigma^2(x))^{-1} \leq 2[4\alpha(1-\alpha)]^{-1}$ .

We now turn to the

Proof of Theorem 4.1. The idea of the proof is as follows. Divide  $I^k$  into a large number of small cubes, and separate these cubes into two groups. In the first group put those cubes over which  $W_F$  has small variance, and use Lemma 4.3 to show that the maximum of  $W_F$  over this group is asymptotically unimportant. For the second group, use Lemma 4.2 to bound the (distribution of the) maximum of  $W_F$  over each cube, then count how many such cubes there are, and thus obtain a final bound.

We now spell out the proof in detail, and note that the only real difficulty lies in finding a convenient labelling system for

the various cubes. We commence with cubes over which  $W_F$  has large variance (i.e., close to  $\frac{1}{2}$ ). Fix the dimension  $k$ , choose  $\alpha \in (\frac{1}{4}, \frac{1}{2})$ ,  $\lambda > 1$ , and set  $\gamma = \alpha^2 / (2k\lambda^2)$ . Let  $\underline{1} = (1, \dots, 1)$ , and let  $x \in I^k$  be such that also  $x + \gamma \underline{1} \in I^k$ . Then the uniformity of the marginals of  $F$  implies

$$(4.17) \quad F(x + \gamma \underline{1}) \leq F(x) + \gamma k = F(x) + \alpha^2 / (2\lambda^2).$$

Now consider the lattice of points of the form  $\gamma(n_1, \dots, n_k)$ , where  $n_i = 0, 1, \dots, [\gamma^{-1}]$ . Then each of these points has a unique expression as  $p + j\gamma \underline{1}$ , where  $p \in \pi$  and  $\pi$  is the set of  $\gamma(n_1, \dots, n_k)$  with  $\min\{n_i : 1 \leq i \leq k\} = 0$ . For each  $p \in \pi$  define, inductively,

$$j_1 = j_1(p) = \max\{j : F(p + j\gamma \underline{1}) < \alpha\},$$

$$j_i = j_i(p) = \max\{j : F(p + j\gamma \underline{1}) - F(p + j_{i-1}\gamma \underline{1}) \leq \alpha^2 / (2\lambda^2)\}.$$

Furthermore, define

$$J = J(p) = \min\{i : F(p + j_i\gamma \underline{1}) > 1 - \alpha\}.$$

Note that (4.17) implies  $j_i - j_{i-1} \geq 1$  for all  $i$  and  $p$ . Also, for  $1 \leq j_1 \leq J$ ,

$$(4.18) \quad 0 < \alpha - \alpha / (2\lambda^2) \leq F(p + j_1\gamma \underline{1}) \leq 1 - \alpha + \alpha^2 / (2\lambda^2) < 1.$$

Now set  $j^*(p) = j_{J(p)} - 1$ , and define

$$S(p) = \bigcup_{k=0}^{j^*(p)} \{x : p + (j_1 + k)\gamma \underline{1} \leq x \leq p + (j_1 + k + 1)\gamma \underline{1}\},$$

and

$$S^*(p) = \bigcup_{i=1}^{J(p)-1} \{x : p + j_i\gamma \underline{1} \leq x \leq p + j_{i+1}\gamma \underline{1}\}.$$

(Drawing a picture for  $k=2$  will undoubtedly make the following argument appear more natural.) From the definitions of  $S$  and  $S^*$  it is clear that

$$\{x: \alpha \leq F(x) \leq 1 - \alpha\} \subseteq \bigcup_{p \in \pi} S(p) \subseteq \bigcup_{p \in \pi} S^*(p) .$$

Thus, with  $A$  as at (4.5)

$$\begin{aligned} (4.18) \quad & P\{\sup(W_F(x): \alpha \leq F(x) \leq 1-\alpha) > \lambda\} \\ & \leq P\{\sup(W_F(x): x \in S^*(p) \text{ for some } p \in \pi) > \lambda\} \\ & \leq \sum_{p \in \pi} \sum_{i=1}^{J(p)-1} P\{A(p+j_i \gamma_1^1, p+j_{i+1} \gamma_1^1, \lambda)\} \\ & \leq \sum_{p \in \pi} O\left(\sum_{i=1}^{J(p)-1} \lambda^{-1} \exp[-\lambda^2/2\sigma^2(p+j_i \gamma_1^1)]\right) , \end{aligned}$$

the last inequality following from Lemma 4.2. Now note that for all  $i, p$ ,

$$F(p + j_{i+2} \gamma_1^1) - F(p + j_i \gamma_1^1) > \alpha^2/2\lambda^2 ,$$

and set  $I = \min\{i: \frac{1}{2} - i\alpha^2/2\lambda^2 < \alpha\} = \min\{i: \frac{1}{2} + i\alpha^2/2\lambda^2 > 1 - \alpha\}$ . Then it follows that the sequence

$$(4.19) \quad \{\sigma^2(p+j_i \gamma_1^1): i=1, \dots, J(p)\}$$

is dominated by the sequence

$$(4.20) \quad \{a_0, a_0, a_0, a_0, a_1, \dots, a_1, a_1, a_1, \dots, a_I, a_I, a_I, a_I\}$$

in which  $a_i = \frac{1}{2} - (i\alpha^2/2\lambda^2)^2$ , where by "domination" we mean that the elements of (4.20) may be rearranged so that, termwise, they dominate corresponding elements of (4.19). Furthermore, there may also be more terms

in (4.20) than in (4.19). As a consequence of this we have that

$$\begin{aligned}
 & \sum_{i=1}^{J(p)-1} \lambda^{-1} \exp[-\lambda^2/2\sigma^2(p+j_i \gamma_1)] \\
 & \leq 4 \sum_{i=1}^I \lambda^{-1} \exp[-\lambda^2/2(\lambda^2 - (i\alpha^2/2\lambda^2)^2)] \\
 & \leq 4 \sum_{i=1}^I \lambda^{-1} \exp[-2\lambda^2(1 + (i\alpha^2/\lambda^2)^2)] \\
 & = 4e^{-2\lambda^2} \sum_{i=1}^I \lambda^{-1} \exp[-2\alpha^4 i^2/\lambda^2] \\
 & = e^{-2\lambda^2} O\left(\int_0^\infty e^{-2\alpha^4 y} dy\right) \\
 & = O(e^{-2\lambda^2}).
 \end{aligned}$$

Note that  $\pi$  has at most  $(2 + 2k\lambda^2/\alpha^2)^{k-1} = O(\lambda^{2(k-1)})$  points. Combining this fact, the above, (4.18) and Lemma 4.3 yields

$$\begin{aligned}
 & P\{\sup_{I^k} W_F(x) > \lambda\} \\
 & \leq P\{\sup(W_F(x): \alpha \leq F(x) \leq 1-\alpha) > \lambda\} + P\{\sup(W_F(x): F(x) < \alpha \\
 & \quad \text{or } F(x) > 1-\alpha) > \lambda\} \\
 & = O(\lambda^{2(k-1)} e^{-2\lambda^2}) + O(e^{-2\beta\lambda^2}) \\
 & = O(\lambda^{2(k-1)} e^{-2\lambda^2}).
 \end{aligned}$$

This completes the proof of Theorem 4.1.

Proof of Theorem 4.2 Our aim here will be to attempt to mimick the proof of the two-dimensional case, Theorem 3.2, by comparing  $W_F$  to  $W_{F_\varepsilon}$ . However,

for  $k > 2$  dimensions the mapping on which the comparison is based is definable as a linear mapping only in an arbitrarily small neighbourhood of the point  $x_0$  of the theorem, and we shall not be able to say anything concrete about the size of the neighbourhood, and thus, a fortiori, anything non-asymptotic about the lower bound that we shall obtain.

The first part of the proof carefully sets up some geometrical structures, and is totally non-probabilistic. Probability will enter only when the groundwork is ready.

Let  $G$  be the uniform distribution on  $I^k$ , and  $y^*$  the point  $(\frac{1}{2})^{1/k}$ . Then  $G(y^*) = \frac{1}{2}$  and

$$(4.21) \quad \gamma := \frac{\partial G}{\partial y_1}(y^*) = (\frac{1}{2})^{(k-1)/k}.$$

Note that  $\gamma$  is independent of  $i$ . In order to compare  $F$  to  $G$ , it is convenient to consider new coordinate systems for  $F$  and  $G$ , obtained by rotation and translation. To this end, let  $\psi = \langle \psi_1, \dots, \psi_k \rangle$ , and write  $\gamma$  interchangeably for the constant (4.21) and the constant vector  $\gamma \mathbf{1}$ . Define the unit vectors, with  $\|\cdot\|$  now denoting the usual Euclidean norm,

$$v_1 = \psi / \|\psi\|, \quad w_1 = \gamma / \|\gamma\|,$$

and extend to two orthonormal bases  $V = \{v_1, \dots, v_k\}$  and  $W = \{w_1, \dots, w_k\}$  for  $R^k$ . Choose the origins of the new spaces to be  $x^*$  and  $y^*$ , respectively. Then if  $v(x)$  and  $w(y)$  are, respectively, the representations of  $x$  and  $y$  in the new coordinate systems, we have

$$v(x^*) = 0, \quad v_1(x) = \psi'(x - x^*) / \|\psi\|$$

$$w(y^*) = 0, \quad w_1(y) = \mathbf{1}'(y - y^*) / \sqrt{k} = \gamma'(y - y^*) / \|\gamma\|.$$

The d.f.'s  $F$  and  $G$  can be transferred in a natural fashion to  $\underline{V}$  and  $\underline{W}$  space, respectively. Let  $\tilde{F}$  and  $\tilde{G}$  be the corresponding functions, defined by

$$\tilde{F}(v(x)) = F(x), \quad \tilde{G}(w(y)) = G(y).$$

(Note that  $\tilde{F}$  and  $\tilde{G}$  are not necessarily d.f.'s on  $\underline{V}$  and  $\underline{W}$  space.) Now define maps  $\pi^V$  and  $\pi^W$  from  $\underline{V}$  and  $\underline{W}$  space, respectively, to the original domains of  $F$  and  $F_\epsilon$ , by

$$\pi^V(v(x)) = x - x^*, \quad \pi^W(w(y)) = y - y^*.$$

Thus,  $\pi^V$  and  $\pi^W$  transform from the coordinate systems of the  $\underline{V}$  and  $\underline{W}$  spaces to systems centered at  $x^*$  and  $y^*$  but oriented like the original cartesian system.

We shall need to impose on the  $\underline{V}$  and  $\underline{W}$  spaces concepts of ordering inherited from the original spaces. To this end, write

$$v^{(1)} < v^{(2)} \iff \pi_1^V(v^{(1)}) \leq \pi_1^V(v^{(2)}), \quad i=1, \dots, k,$$

$$w^{(1)} < w^{(2)} \iff \pi_1^W(w^{(1)}) \leq \pi_1^W(w^{(2)}), \quad i=1, \dots, k,$$

and define  $v^{(1)} \hat{<} v^{(2)}$  and  $w^{(1)} \hat{<} w^{(2)}$  accordingly.

This completes the necessary geometrical groundwork. We now build the mapping upon which the comparison between  $F$  and  $G$  will be based. Let

$$\rho = \frac{||\gamma||}{||\psi||} = \frac{\sqrt{\rho} \gamma}{||\psi||}, \quad \alpha = \frac{p \max\{\psi_i : i=1, \dots, k\}}{\min\{\psi_i : i=1, \dots, k\}}, \quad \beta = \rho + 2\alpha.$$

Define the mapping  $m = (m_1, \dots, m_k)$  from a neighbourhood of zero in  $\underline{W}$  space to a neighbourhood of zero in  $\underline{V}$  space, via its coordinate mappings, by firstly setting

$$(4.22) \quad m_i(w) = \beta w_i, \quad i=2, \dots, k$$

and then choosing  $m_1(w)$  such that

$$(4.23) \quad \tilde{F}(m(w)) = \tilde{G}(w) .$$

We need to check that  $m_1$  is, in fact, well defined. For  $w_1 = 0$ ,  $i=2, \dots, k$ , and general  $w_1$ ,  $\tilde{G}(w)$  is clearly strictly increasing as a function of  $w_1$ . Furthermore, since the unit vector  $V_1$  of the  $V$  space has, as a vector in the original space, strictly positive coordinates, it follows that  $\tilde{F}(V_1, 0, \dots, 0)$  is strictly increasing as a function of  $V_1$ . Since  $\tilde{F}(0) = \tilde{G}(0) = \frac{1}{2}$ , it follows that  $m$  is well defined for  $w$  of the form  $(w_1, 0, \dots, 0)$ . The implicit function theorem now defines  $m$  uniquely for sufficiently small neighbourhoods.

Having defined our mapping, let us consider some of its properties. Note that for small neighbourhoods of the origin

$$\tilde{G}(w) = \frac{1}{2} + ||\gamma||w_1 + o(||w||) ,$$

$$\tilde{F}(v) = \frac{1}{2} + ||\psi||v_1 + o(||v||) .$$

Combining these facts with (4.22), (4.23), and the definition of  $\beta$  we obtain that for small  $w$

$$(4.24) \quad m_1(w) = \rho w_1 + o(||w||) .$$

Consequently, for small  $w^1$ ,

$$(4.25) \quad |[m_1(w^1) - m_1(w^2)] - [\rho w_1^1 - \rho w_1^2]| \leq ||w^1 - w^2||$$

Now let  $q$  be the linear map approximating  $m$ ; i.e., set

$$(4.26) \quad q_1(w) = \rho w_1, \quad q_i(w) = \beta w_i, \quad i=2, \dots, k .$$



Then by (4.25), for small  $w^1$ ,

$$(4.27) \quad |[m(w^1) - m(w^2)] - [q(w^1) - q(w^2)]| \leq \|w^1 - w^2\|.$$

Finally, note that as a consequence of (4.28) we also have

$$(4.2) \quad \pi_1(m(w^1) - m(w^2) - (q(w^1) - q(w^2))) \leq \|w^1 - w^2\|.$$

This completes our listing of properties of  $m$  and its linear approximation. We can now turn to the final part of the proof, the comparison of  $W_F$  and  $W_G$ , which we commence by comparing  $\tilde{F}$  and  $\tilde{G}$ .

Firstly, let  $N$  be a small enough neighbourhood of zero in  $W$  space so that (4.24) - (4.28) are true for  $w^1 \in N$ . Take  $w^1, w^2 \in N$  with  $w^1 \wedge w^2 \in N$ . Suppose

$$(4.29) \quad w^1 \wedge w^2 = w^p \quad \text{for } p = 1 \text{ or } 2.$$

Then

$$\begin{aligned} (4.30) \quad \tilde{F}(m(w^1) \wedge m(w^2)) &= \tilde{F}(m(w^1)) \wedge \tilde{F}(m(w^2)) \\ &= \tilde{G}(w^1) \wedge \tilde{G}(w^2) \\ &= \tilde{G}(w^p) \\ &= \tilde{G}(w^1 \wedge w^2) \end{aligned}$$

Now consider the case  $w^3 = w^1 \wedge w^2 \neq w^p$  for either  $p = 1$  or  $2$ . We shall obtain (4.30) also for this case, but with inequality replacing the equality. For each coordinate  $j=1, \dots, k$ ,  $w^3 = w^1 \wedge w^2$  implies that

$$\pi_j^w(w^1) - \pi_j^w(w^3) = 0 \quad \text{or} \quad \pi_j^w(w^2) - \pi_j^w(w^3) = 0.$$

Fix  $j$ , and let  $p := p(j) = 1$  or  $2$  be such that

$$(4.31) \quad \pi_j^w(w^p) - \pi_j^w(w^3) = 0.$$

Thus  $w^3 = \langle w_1^{p(1)}, \dots, w_k^{p(k)} \rangle$ . Rewrite  $q$  as

$$q(w) = \beta w + (\rho - \beta) w_1 v_1,$$

and apply this to (4.31) with  $w = w^p - w^3$  to obtain, via the linearity of  $\pi_j$ ,

$$(4.32) \quad \pi_j^w(q(w^p - w^3)) = -2\alpha(w_1^p - w_1^3) \pi_j^w(v_1).$$

Now note, from the definition of  $w_1$  and since  $w^m \gg w^3$ ,

$$w_1^p - w_1^3 = \sum_{i=1}^k \pi_i^w(w_1^p \cdot w_1^3) / \sqrt{k} \geq ||w^p - w^3|| / \sqrt{k}$$

Furthermore, it follows from the definition of  $v_1$  and  $\alpha$  that

$$\pi_j^w(v_1) \geq \sqrt{k} \alpha.$$

Substituting the above two inequalities into (4.32) yields

$$\pi_j^w(q(w^p - w^3)) \leq -2||w^p - w^3||.$$

Combining this with (4.28) thus yields

$$(4.33) \quad \pi_j^w(m(w^p) - m(w^3)) \leq -||w^p - w^3|| < 0.$$

However, what we have just shown is that for every  $j = 1, \dots, k$  there is a  $p = p(j)$  satisfying (4.33). Consequently, for every  $w^1, w^2 \in N$  with  $w^1 \hat{=} w^2 \in N$ , it follows that

$$m(w^1 \hat{=} w^2) \gg (m(w^1) \hat{=} m(w^2)),$$

from which it follows that

$$(4.34) \quad \begin{aligned} \tilde{F}(m(w^1) \hat{=} m(w^2)) &\leq \tilde{F}(m(w^1 \hat{=} w^2)) \\ &= \tilde{G}(w^1 \hat{=} w^2). \end{aligned}$$

Combining this with (4.30) we find that the above inequality holds for all  $w^1, w^2 \in N'$ , where  $N' \subset N$  is a neighbourhood of zero such that  $w^1, w^2 \in N'$  implies  $w^1 \wedge w^2 \in N$ , and, consequently, that (4.34) holds.

To obtain the final comparison between  $F$  and  $G$ , we need to return to the original coordinate system. However, this is now easy, for since the "minimum" relationship in (4.34) is really that of the original coordinate system it trivially follows that via  $m$  we have constructed a map, say  $m^*$ , from some neighbourhood  $N^*$  of  $y^*$  to a neighbourhood  $m^*(N^*)$  of  $x^*$  satisfying

$$F(m^*(y)) = G(y), \quad y \in N^*,$$

$$F(m^*(y^1) \wedge m^*(y^2)) = G(y^1 \wedge y^2), \quad y^1, y^2 \in N^*.$$

Slepian's inequality now yields

$$\begin{aligned} (4.35) \quad P\{\sup_{I^k} W_F(x) > \lambda\} &\geq P\{\sup_{m(N^*)} W_F(x) > \lambda\} \\ &\geq P\{\sup_{N^*} W_G(y) > \lambda\} \\ &\geq P\{\sup_{B_\delta} W_G(y) > \lambda\}, \end{aligned}$$

where  $\delta$  is chosen small enough so that  $B_\delta := [(\frac{1}{2})^{1/k-\delta}, (\frac{1}{2})^{1/k+\delta}]^k \subset N^*$ .

(Note that  $\delta$  depends on  $N^*$ , and so on  $F$ .) Thus, to complete the proof, we need only find a lower bound for the last probability.

Take  $\epsilon = \frac{1}{2}(2\delta)^k$ , and consider the d.f.  $F_\epsilon$  of Section 2 on  $A_\epsilon$ .

Map  $y \in B_\delta$  to  $z \in A_\epsilon$  according to the coordinate mappings

$$z_1 = z_1(y) = \frac{1}{2} + (2\delta)^{k-1} [y_1 - (\frac{1}{2})^{1/k}].$$

Then it is straightforward to check that for  $y \in B_\delta$

$$(4.36) \quad F_\varepsilon(z(y)) = G(y) .$$

Furthermore, if  $y^1, y^2 \in B_\delta$  then  $y^1 \wedge y^2 \in B_\delta$ , and

$$(4.37) \quad z(y^1 \wedge y^2) = z(y^1) \wedge z(y^2) .$$

Consequently,

$$P\{\sup_{B_\delta} W_G(y) > \lambda\} = P\{\sup_{A_\varepsilon} F_\varepsilon(z) > \lambda\}.$$

But this last probability is known, and is bounded from below in Theorem 2.2 by  $c\lambda^{2(k-1)}e^{-2\lambda^2}$ . Combining this with (4.36) completes the proof of the theorem.

### 5. AN UPPER-LOWER CLASS THEOREM

We now return to the one-sided KS statistic  $T_n^{(k)}$  of the introduction, and study the way it grows with  $n$ . In a fundamental paper treating the one-dimensional case Chung (1949) proved the following result for a sequence  $\lambda(n) \uparrow \infty$ :

$$(5.1) \quad P\{T_n^{(1)} > \lambda(n) \text{ infinitely often}\} = 0(1)$$

if

$$(5.2) \quad \sum_n \frac{\lambda^2(n)}{n} e^{-2\lambda^2(n)} < \infty \quad (= \infty).$$

Kiefer (1961) obtained a weaker version of Chung's result for the multivariate case, and proved the following LIL for every  $k$  and continuous  $F$ :

$$(5.3) \quad 1 = P\{\limsup_{n \rightarrow \infty} \frac{T_n^{(k)}}{(\frac{1}{2} \log \log n)^{\frac{1}{2}}} = 1\} = P\{\liminf_{n \rightarrow \infty} \frac{T_n^{(k)}}{(\frac{1}{2} \log \log n)^{\frac{1}{2}}} = -1\}.$$

Kiefer's proof of (5.3) was based on inequality (1.3), which is not fine enough to pick up the higher iterated logarithm terms that (5.2) yields. Having improved on Kiefer's inequality in the previous sections (at least insofar as the limit process  $W_F$  is concerned) we can now complete the task Kiefer began and obtain a multi-dimensional analogue of (5.1).

Unlike Chung's and Kiefer's basic inequalities for  $P\{T_n > \lambda\}$ , we have only inequalities for  $P\{\sup_x W_F(x) > \lambda\}$ , and so we shall need to proceed via an embedding theorem. To this end, for continuous  $F$  on  $I^k$  define the Kiefer process as the  $C[0,1]^k$ -valued, real parameter process  $K_t$ ,  $t \geq 0$ , satisfying:

$$(5.4) \quad P\{K_1 \in A\} = P\{W_F \in A\}.$$

$$(5.5) \quad P\{(K_t - K_s) \in A\} = P\{K_{t-s} \in A\} = P\{\sqrt{t-s} K_1 \in A\}, \text{ for } t > s,$$

$$(5.6) \quad (K_t - K_s) \text{ and } K_u \text{ are independent for all } t > s \geq u.$$

Here  $A$  is any Borel subset of  $C[0,1]^k$ , with topology generated by the sup norm  $\|k\| = \sup\{|k(t)| : t \in I^k\}$ . Then Theorem 7.1 of Dudley and Philipp (1983) implies the following embedding theorem, which is a strengthening of an earlier result of Kiefer (1972).

Theorem 5.1 (Dudley-Phillip) Let  $X_1, X_2, \dots$  be an infinite sequence of i.i.d.r.v.'s, defined on an infinite product space  $(R^\infty, B^\infty, P^\infty)$  with common d.f.  $F$ . Let  $(\Omega, \Sigma, Pr)$  be the product of  $(R^\infty, B^\infty, P^\infty)$  and a copy of the unit interval with Lebesgue measure. Let  $F_n$  be the empirical d.f. based on  $X_1, \dots, X_n$ . Then, for every  $\theta > 0$  there exists a Kiefer process  $K_t, t \geq 0$ , defined on  $\Omega$ , such that

$$(5.7) \quad \sup_x |n[F_n(x) - F(x)] - K_n(x)| \leq O(n^{1/2}(\log n)^{-\theta})$$

with  $Pr$  probability one.

As an immediate consequence of this result, along with a LIL for sums of Banach space random variables, it is now easy to obtain Kiefer's LIL, (5.3). (c.f. Kuelbs and Philipp (1980) and Goodman, Kuelbs and Zinn (1981), esp. Theorem 6.1). Indeed, the Banach space results yield much more than (5.3), for they also identify the cluster points of  $F_n$  in  $C[0,1]^k$  in terms of the unit ball of a certain Hilbert space. It is not possible, however, to follow this path to obtain a multivariate

version of Chung's upper-lower class theorem, the problem being that no appropriate upper-lower class theorem is known for  $K_t$ . (Note that whereas Kuelbs (1975) does have a result of this type for  $K_t$ , it is not applicable here, since it gives results not for the growth of  $\|K_t\|$  but the growth of  $\|K_t\|_*$ , where  $\|\cdot\|_*$  is another unspecified norm (albeit equivalent to the sup norm).) Consequently we shall have to revert to an almost basic principles analysis to obtain a generalization of Chung's theorem.

To state our result, we shall say a non-negative, non-decreasing, continuous function  $\psi(t)$  defined for large values of  $t$  is a lower function for  $\{K_t, t \geq 0\}$  if

$$(5.8) \quad P\{\|K_n\| > \sqrt{n} \psi(n) \text{ for an unbounded set of } n\text{'s}\} = 1,$$

and an upper function for  $\{K_t: t \geq 0\}$  if

$$(5.9) \quad P\{\|K_t\| > t^{\frac{1}{2}} \psi(t) \text{ for only a bounded set of } t\text{'s}\} = 1.$$

Since the definition of  $K_t$  is dependent on  $F$ , whether or not any given  $\psi$  is a lower or upper function depends on  $F$  as well as  $\psi$ . Thus we write  $\psi \in L(F)$  and  $\psi \in U(F)$ , respectively, to denote this dependency. Note that (5.8) implies the weaker condition,

$$P\{\|K_t\| > t^{\frac{1}{2}} \psi(t) \text{ for an unbounded set of } t\text{'s}\} = 1$$

which is usually taken as the definition of a lower function. However, the stronger result (5.8) is what is needed to apply Theorem (5.1), and since our proof will be strong enough to prove (5.8) we use it to define the notion of lower class. We can now state

Theorem 5.2 Let  $F$  be continuous on  $I^k$  with uniform marginals. For  $\psi$  as above, set

$$(5.10) \quad I_k(\psi) = \int \frac{\psi^{2k}(t)}{t} e^{-2\psi^2(t)} dt.$$

If  $I_k(\psi) < \infty$ , then  $\psi \in \mathcal{U}(F)$ . Furthermore, if  $F$  satisfies the conditions of Theorem 4.2, and  $I_k(\psi) = \infty$ , then  $\psi \in \mathcal{L}(F)$ .

A simple argument, dating back at least to Erdos (1942) and spelled out in detail in Sirao (1959), shows that there is no loss of generality in Theorem 5.2 in assuming that for large  $t$

$$(5.11) \quad (\frac{1}{2} \log \log t)^{\frac{1}{2}} \leq \psi(t) \leq (\log \log t)^{\frac{1}{2}}.$$

Furthermore, a straightforward application of the Abel-Dini theorem easily yields the following corollary.

Corollary 5.1 Let  $p \geq 3$  be integral, and define

$$\psi_{k,\delta}(t) = 2^{-\frac{1}{2}} [\log_2 t + (k+1) \log_3 t + \log_4 t + \dots + (1+\delta) \log_{p+1}(t)]^{\frac{1}{2}}.$$

Then  $\delta > 0$  implies  $I_k(\psi_{k,\delta}) < \infty$  and  $\delta < 0$  implies  $I_k(\psi_{k,\delta}) = \infty$ , so that  $\psi_{k,\delta} \in \mathcal{U}(F)$  if  $\delta > 0$  and  $\psi_{k,\delta} \in \mathcal{L}(F)$  if  $\delta < 0$  and  $F$  satisfies the conditions of Theorem 4.2.

As a further consequence of (5.11) and Theorem 5.1 we can also derive the following corollary of Theorem 5.2, which generalises Chung's uni-variate test:



Corollary 5.2 For all  $F$ ,

$$P\{T_n^{(k)} > \psi(n) \text{ i.o.}\} = 0 \text{ if } \sum_n \frac{\psi^{2k}(n)}{n} e^{-2\psi^2(n)} < \infty$$

For  $F$  satisfying the conditions of Theorem 4.2

$$P\{T_n^{(k)} > \psi(n) \text{ i.o.}\} = 1 \text{ if } \sum_n \frac{\psi^{2k}(n)}{n} e^{-2\psi^2(n)} = \infty.$$

This result, of course, implies Kiefer's LIL, (5.3). All that now remains is the

Proof of Theorem 5.2. We consider the convergent case first, i.e.,  $I_k(\psi) < \infty$ .

Define a sequence  $t_n$  satisfying

$$(5.12) \quad t_{n+1} = t_n(1 + \psi^{-2}(t_n)),$$

where  $t_1 > 3$  is sufficiently large so that (5.11) holds for  $t > t_1$ , and so  $\lim_n t_n = \infty$ . Set  $I_n = [t_n, t_{n+1}]$  and

$$A_n = \left\{ \sup_{t \in I_n} \frac{||K_t||}{t^{\frac{1}{2}\psi(t)}} > 1 \right\}.$$

Then, applying the Banach space version of Lévy's inequality, we have

$$\begin{aligned} P\{A_n\} &\leq P\left\{\sup_{t \in I_n} ||K_t|| > t_n^{\frac{1}{2}\psi(t_n)}\right\} \\ &\leq 2P\{||K_{t_{n+1}}|| > t_n^{\frac{1}{2}\psi(t_n)}\} \\ &= 2P\{t_{n+1}^{-\frac{1}{2}} ||K_{t_{n+1}}|| > (t_n/t_{n+1})^{\frac{1}{2}\psi(t_n)}\}. \end{aligned}$$

Now apply the scaling law (5.5), and Theorem 4.1 to obtain

$$P\{A_n\} \leq C(t_n/t_{n+1})^{k-1} [\psi(t_n)]^{2(k-1)} \exp[-2\psi^2(t_n)t_n/t_{n+1}]$$

$$\leq C[\psi(t_n)]^{2(k-1)} \exp[-2\psi^2(t_n)] ,$$

since  $t_n < t_{n+1}$  and  $(t_n/t_{n+1}) = (1+\psi^{-2}(t_n))^{-1} > 1-\psi^{-2}(t_n) > \frac{1}{2}$  for large enough  $n$ .

To complete the proof it is clearly sufficient to show  $\sum_{n=1}^{\infty} P\{A_n\}$  converges. But

$$\sum_{n=1}^{\infty} P\{A_n\} \leq C \sum_n \int_{t_{n-1}}^{t_n} [\psi(t_n)]^{2(k-1)} e^{-2\psi^2(t_n)} \frac{t_n}{s(t_n - t_{n-1})} ds$$

$$\leq C \sum_n \int_{t_{n-1}}^{t_n} [\psi(t_n)]^{2(k-1)} e^{-\psi^2(t_n)} \frac{t_n \psi^2(t_{n-1})}{s t_{n-1}} ds$$

$$\leq C \int_0^{\infty} \psi^{2k}(s) e^{-2\psi^2(s)} ds$$

$$= I_k(\psi) ,$$

the last inequality from the definition of  $t_k$  and the ultimate monotonicity of  $\psi^{2k}(s)e^{-2\psi^2(s)}$ . By assumption  $I_k(\psi) < \infty$ , and so the proof of the convergent part of the theorem is complete.

Now assume  $I_k(\psi) = \infty$ . Let  $\alpha_n = (\log n)^2$ ,

$$\beta_n = \sum_{i=2}^n \alpha_i , \text{ and } t_n = [\beta n] , n \geq 2 . \text{ Also set}$$

$$\epsilon_n = -1 + \frac{(1 + t_n/t_{n+1})}{(1 - t_n/t_{n+1})^{1/2}} > 0 , n \geq 2 .$$

Following Chung's (1949) argument, set

$$H_n = \{ ||K_{t_n}|| < t_n^{\frac{1}{2}} \psi(t_n) \},$$

$$H_{n,n+1} = \{ ||K_{t_{n+1}} - K_{t_n}|| > (1+\epsilon_n)(t_{n+1}-t_n)^{\frac{1}{2}} \psi(t_{n+1}) \}.$$

Then, if both  $H_n$  and  $H_{n,n+1}$  occur, we have

$$\begin{aligned} ||K_{t_{n+1}}|| &\geq (1+\epsilon_n)(t_{n+1}-t_n)^{\frac{1}{2}} \psi(t_{n+1}) - t_n^{\frac{1}{2}} \psi(t_n) \\ &= t_{n+1}^{\frac{1}{2}} \psi(t_{n+1}) \{ (1-t_n/t_{n+1})^{\frac{1}{2}} (1+\epsilon_n) - t_n^{\frac{1}{2}} \psi(t_n) / [t_{n+1}^{\frac{1}{2}} \psi(t_{n+1})] \} \\ &\geq t_{n+1}^{\frac{1}{2}} \psi(t_{n+1}). \end{aligned}$$

That is,  $H_n \cdot H_{n,n+1} \Rightarrow \bar{H}_{n+1}$ . Thus, noting from (5.6) that  $H_{n,n+1}$  and  $H_m$  are independent for  $m \leq n$ , it follows that

$$P\{ \prod_{m=2}^n H_m \} \cdot P\{ H_{n,n+1} \} \leq P\{ \prod_{m=2}^n H_m \cdot \bar{H}_{n+1} \}.$$

That is

$$\begin{aligned} P\{ \prod_{m=2}^{n+1} H_m \} &\leq P\{ \prod_{m=2}^n H_m \} \cdot (1-P\{ H_{n,n+1} \}) \\ &\leq P\{ H_2 \} \cdot \prod_{m=2}^n (1-P\{ H_{m,m+1} \}) \end{aligned}$$

Clearly, then, if we can show  $\sum_{n=2}^{\infty} P\{ H_{n,n+1} \} = \infty$ , we will have proven the theorem. Applying Theorem 4.2 and (5.5) we have

$$(5.13) \quad P\{H_{n,n+1}\} \geq C(1+\epsilon_n)^{2(k-1)} [\psi(t_{n+1})]^{2(k-1)} \exp\{-2\psi^2(t_{n+1})(1+\epsilon_n)^2\}.$$

Consider the exponent, and note that

$$\frac{1}{\alpha_{n+1}} - \frac{1}{\beta_{n+1}} < \frac{t_n}{t_{n+1}} < \frac{1}{\alpha_{n+1} - \beta_n^{-1}} < \frac{1}{\alpha_{n+1}} + O(1/\beta_n \alpha_{n+1}).$$

Consequently

$$(1+\epsilon_n)^2 = 1 + 3/\alpha_{n+1} + O(\alpha_{n+1}^{-2}),$$

so that, by (5.11),

$$\begin{aligned} \psi^2(t_{n+1})(1+\epsilon_n)^2 &\leq \psi^2(t_{n+1}) + (\log \log \beta_{n+1}) \cdot 4/\alpha_{n+1} \\ &= \psi^2(t_{n+1}) + \frac{\log \sum_{m=2}^{n+1} \log^2 m}{4 \log^2(n+1)} \\ &\leq \psi^2(t_{n+1}) + 1, \end{aligned}$$

for large enough  $n$ . Substituting into (5.13) and setting  $\gamma_n = (\alpha_{n+1})^{1+1/n}$  yields

$$\begin{aligned} \mathbb{E}P\{H_{n,n+1}\} &\geq C \sum_{n=3}^{\infty} [\psi(t_n)]^{2(k-1)} e^{-2\psi^2(t_n)} \\ &\geq C \sum_n \int_{\alpha_n}^{\gamma_n} [\psi(t_n)]^{2(k-1)} e^{-2\psi^2(t_n)} \frac{\alpha_n}{s(\gamma_n - \alpha_n)} ds \\ &\geq C \sum_n \int_{\alpha_n}^{\alpha_{n+1}} \frac{[\psi(s^n)]^{2k} e^{-2\psi^2(s^n)}}{s} \cdot \frac{\alpha_n}{\psi^2(\beta_n)(\gamma_n - \alpha_n)} ds \end{aligned}$$

since  $\psi^{2k}(t)e^{-2\psi^2(t)}$  is eventually decreasing in  $t$ . A change of variables leads to

$$\sum_n P\{H_{n,n+1}\} \geq C \sum_n a_n \frac{(\alpha_{n+1})^{n+1}}{(\alpha_n)^n} \frac{[\psi(t)]^{2k} e^{-2\psi^2(t)}}{t} dt,$$

where

$$(5.14) \quad a_n = \frac{\alpha_n}{n\psi^2(\beta_n)(\gamma_n - \alpha_n)}.$$

If we can now show that  $a_n$  is bounded away from zero for large enough  $n$ , we shall have  $\sum P\{H_{n,n+1}\} \geq C I_k(\psi) = \infty$ , and the proof will be complete.

Firstly, note that by (5.11)

$$(5.15) \quad \psi^2(\beta_n) \leq \log \log \left( \sum_{m=2}^n \log^2 m \right) = \log \left( \sum_{m=2}^n \log (\log^2 m) \right) \\ \leq \log ((n-2)2 \log \log n) \\ \leq 2 \log n.$$

Furthermore

$$(5.16) \quad \alpha_n / (\gamma_n - \alpha_n) > \alpha_n / (\alpha_{n+1} - \alpha_n) = \left[ \frac{(\log n+1)^2}{(\log n)^2} - 1 \right]^{-1} \\ \geq \frac{1}{2} \left[ \frac{\log n+1}{\log n} - 1 \right]^{-1}.$$

Substituting (5.15) and (5.16) into (5.14) yields

$$a_n \geq \frac{1}{2} [n(\log(n+1) - \log n)]^{-1} \\ = \frac{1}{2} [n \log (1 + 1/n)]^{-1} \\ > 1/8,$$

which completes the proof of the theorem.

# REFERENCES

- [1] ALEXANDER, K. (1983) Probability inequalities for empirical processes and a law of the iterated logarithm. Reprint.
- [2] BROWN, L.D. and ADLER, R.J. (1984) Multivariate Kolmogorov-Smirnov tests, in preparation.
- [3] CABAÑA, E.M. (1982) On the transition density of multidimensional parameter Wiener process with one barrier. Technical Report 78, Departamento de Matematicas y Ciencia de la Computacion, Univ. Simon Bolivar, Venezuela.
- [4] CABAÑA, E.M. and WSCHBOR, M. (1982). The two parameter Brownian bridge: Kolmogorov inequalities and upper and lower bounds for the distribution of the maximum. Ann. Probability. 10 289-302.
- [5] CHUNG, K.L. (1949). An estimate concerning the Kolmogoroff limit distribution. Trans. Amer. Math. Soc. 67 36-50.
- [6] DONSKER, M. (1951) An invariance principle for certain probability limit theorems. Mem. Amer. Math. Soc. 6.
- [7] DONSKER, M. (1952) Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. Ann. Math. Statis. 23 277-281.
- [8] DOOB, J.L. (1949) Heuristic approach to the Kolmogorov-Smirnov theorems. Ann. Math. Statist. 20 393-403.
- [9] DUDLEY, R.M. (1966). Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. Illinois J. Math. 10 109-126.
- [10] DUDLEY, R.M. (1967) Measures on non-separable metric spaces. Illinois J. Math. 11 449-453.

- [11] DUDLEY, R.M. (1978) Central limit theorems for empirical measures. Ann. Probability 6 899-929.
- [12] DUDLEY, R.M. and PHILIPP, W. (1983) Invariance principles for sums of Banach space valued random elements and empirical processes. Z. Wahrscheinlichkeitstheorie verw. Gebiete. 62 509-552.
- [13] ERDOS, P. (1942) On the law of the iterated logarithm. Ann. Math. 43 419-436.
- [14] FELLER, W. (1948) On the Kolmogorov-Smirnov limit theorems for empirical distributions. Ann. Math. Statist. 19 177-189.
- [15] FELLER, W. (1971) An Introduction to Probability Theory and its Applications, Vol. 2, Wiley, N.Y.
- [16] FERNIQUE, X. (1970) Intégrabilité des vecteurs gaussiens. C.R. Acad. Sci. Paris Ser. A. 270 1698-1699.
- [17] FERNIQUE, X. (1971) Régularité des processus gaussiens. Invent. Math. 12 304-320.
- [18] GOODMAN, V. (1976) Distribution estimates for functionals of the two parameter Wiener process. Ann. Probability. 4 977-982.
- [19] GOODMAN, V., KUELBS, J. and ZINN, J. (1981) Some results on the LIL in Banach space with applications to weighted empirical processes. Ann. Probability. 9 713-752.
- [20] KAC, M., KIEFER, J. and WOLFOVITZ, J. (1955). On tests of normality and other tests of goodness of fit based on distance methods. Ann. Math. Statist. 26 189-211.
- [21] KIEFER, J. (1961) On large deviations of the empiric distribution function of vector chance variables and a law of the iterated logarithm. Pacific J. Math. 11 649-660.

- [22] KIEFER, J. (1972) Skorohod embedding of multivariate R.V.'s and the sample DF. Z. Wahrscheinlichkeitstheorie verw. Gebiete. 24 1-35.
- [23] KIEFER, J. and WOLFOVITZ, J. (1958). On the deviations of the empiric distribution function of vector chance variables. Trans. Amer. Math. Soc. 87 173-186.
- [24] KOLMOGOROV, A.N. (1933). Sulla determinazione empirica di una legge di distribuzione. Inst. Ital. Atti. Giorn. 4 83-91.
- [25] KUELBS, J. (1975) Sample path behaviour for Brownian motion in Banach spaces. Ann. Probability 3 247-261.
- [26] KUELBS, J. and PHILIPP, W. (1980) Almost sure invariance principles for partial sums of mixing B-valued random variables. Ann. Probability 8 1003-1036.
- [27] LANDAU, H.J. and SHEPP, L.A. (1971) On the supremum of a Gaussian process. Sankya Ser. A 32 369-378.
- [28] LEADBETTER, M.R., LINDGREN, G., And ROOTZEN, H. (1983) Extremes and Related Properties of Random Sequences and Processes. Springer, New York.
- [29] MARCUS, M.B. and SHEPP, L.A. (1971) Sample behaviour of Gaussian processes. Proc. Sixth Berkeley Symp. Math. Statist. and Probability II. University of California Press 423-441.
- [30] PARK, C. and SKOUG, D.L. (1978) Distribution estimates of barrier-crossings probabilities of the Yeh-Wiener process. Pacific J. Math. 78 455-466.
- [31] SIRAO, T. (1959) On the continuity of Brownian motion with a multidimensional parameter. Nagoya Math. J. 16 135-156.
- [32] SLEPIAN, D. (1962) The one-sided barrier problem for Gaussian noise. Bell System Tech. J. 41 463-501.
- [33] SMIRNOV, N.V. (1944) Approximation of the distribution law of a random variable by empirical data (In Russian). Uspehi Matematicheski Nauk.



END

FILMED

12-84

DTIC